Arkwood, Inc., Superfund Site
Comments on Revised Conceptual Site Model

Item No.	Reference	Comments by EPA Dated July 18, 2014	PRP Response Dated: August 29, 2014
1.	Revised CSM General	EPA has updated the Standard Default Exposure Factors used by the Superfund program: http://www.epa.gov/oswer/riskassessment/pdf/superfund-hh-exposure/OSWER-Directive-9200-1-120-ExposureFactors.pdf These Standard Default Exposure Parameters are used in the Regional Screening Level (RSL) calculator. http://www.epa.gov/region6/6pd/rcra_c/pd-n/screen.htm The RSL calculator shows, as of May 2014, the updated non-cancer screening level for 2,3,7,8-TCDD toxicity equivalent quotient (TEQ) for an industrial worker is 730 part per trillion (ppt).	
2.	Revised CSM Table 1 ROD p. 64-65	The Record of Decision (ROD) originally set the dioxin clean-up level at 20,000 ppt, and it required 6-12 inches of clean soil as a cover. Data in Table 1 of the Revised CSM indicate there are many areas where dioxin concentrations in soil under the cover exceed the new screening level of 730 ppt. However, EPA understands that many of the additional samples identified in the Revised CSM are to be collected from the cover soil, not from the native soil below it. Therefore, concerns about the level of heterogeneity in the soil to be tested may or may not be applicable at this point, but they are presented below for informational purposes.	
3.	Revised CSM page 1 ICS User Guide Page 29 Item 11	The CSM states that "the USEPA (2011) guidance for incremental composite soil sampling was utilized…." However, some proposed actions do not follow the guidance. One of these is the approach for determining compliance with the TEQ screening level. The CSM uses the term decision unit (DU) presumably using	

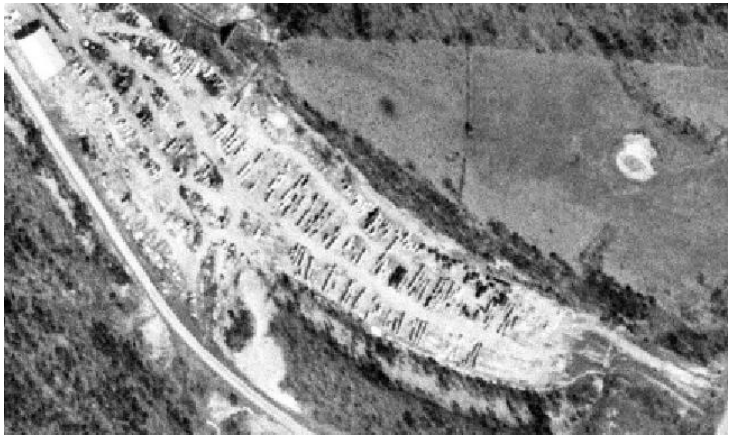
		the term as the ICS User Guide does. The definition of DU according to the ICS guidance is "the volume of soil over which *a mean concentration value is obtained for comparison to a regulatory threshold val*ue." [emphasis added] Yet the proposed plan says that "the maximum composite measurement for each DU [will be compared] to the dioxin soil screening level of 665 ppt TEQ." (Note that the screening level has now changed to 730 ppt. See comment No. 1.) This proposed approach conflicts with the ICS guidance. The ICS User Guide relies on statistical calculations to determine an estimate of the average concentration for a decision unit (DU). By using the 95% upper confidence interval (UCL) as the statistical estimate of the DU mean, the likelihood of making a decision error about whether or not the true DU average exceeds the screening level is controlled to 5% or less.	
4.	Revised CSM General	Generating a UCL for the average DU concentration requires having an estimate of variability. There are at least 3 options for doing this. In this comment EPA presents 2 that are relevant here. 1) When an incremental sample covers an entire DU, the estimate of variability is usually obtained by collecting at least 3 independent replicate DU-ICS samples (i.e., 3 separate ICS samples, each having 30 or more increments, and each of which evenly covers the entire DU area). Since each DU-ICS sample represents an average for the DU, this provides 3 independent estimates of the DU mean. a. If the 3 DU-ICS replicate samples agree within 20% RSD, the 3 DU-ICS values can be used to calculate a Student's-t UCL. b. If the RSD for the 3 DU-ICS replicate samples exceeds 20% RSD, a Chebyshev UCL should be calculated. If the Chebyshev UCL exceeds the decision threshold, i. The ICS replication QC/variability partitioning results should be evaluated to identify effective	

		corrective actions to the sampling design, or ii. The decision may be that the DU is assumed to exceed the threshold. 2) If the DU is divided into subunits (which are called sampling units (SUs), which is what this CSM proposes, the variability between the SU results can be used in the calculation of the DU's UCL. **Note that a few things are different from scenario 1 above.** a. SU-ICS samples are NOT replicates because they do not cover the same soil volume. b. Individual SU-ICS sample are not representative of the entire DU because they do not cover the entire DU. c. Either all of the SUs comprising a DU must be sampled, or, if there are enough SUs in a large DU, a statistically valid subset of the SUs can be sampled. Since SUs cover different portions of the DU, a normal distribution of SU-IS results cannot be expected. Therefore, the following guidelines are suggested to avoid the time and expense of follow-up sampling events to address data uncertainties: i. If the SUs to be sampled are selected randomly, enough SU-IS results must be available to determine the distribution of SU-IS results (so that the correct type of parametric or nonparametric UCL can be selected). 1. If the average DU concentration is expected to be near the action level/decision threshold… a. …and the DU is expected to be fairly homogeneous within its borders (i.e., SUs will probably have similar results), no less than 7 SU-IS results are needed. b. …and the DU is expected to be heterogeneous within its borders (i.e., SUs will probably have very different	

		results), no fewer than 10 SU-IS results are needed. c. If the number of SUs in the DU is fewer than these limits, then all SUs should be sampled. 2. If the average DU concentration is expected to be far above or below the action level/decision threshold… a. …and the DU is expected to be fairly homogeneous within its borders (i.e., SUs will probably have similar results), no less than 5 SU-IS results are needed. b. …and the DU is expected to be heterogeneous within its borders (i.e., SUs will probably have very different results), no fewer than 7 SU-IS results are needed. c. If the number of SUs in the DU is fewer than these limits, then all SUs should be sampled. ii. If the number of SU-ISs actually collected is less than these guidelines, the nonparametric Chebyshev UCL must be used since there is likely insufficient data for reliable results from parametric distribution tests. 1. If the Chebyshev UCL exceeds the decision threshold: a. Either the DU must be considered to exceed the threshold, or b. Additional SU-ISs can sampled to add to the data set so that the UCL might be calculated as less than the threshold. 2. If the Chebyshev UCL does not exceed the decision threshold: a. Evaluate the actual between-SU variability for this DU AND in	

		neighboring or similar DUs for indications that the sampling design was inadequate to capture the full range of variability. b. Determine the highest between-SU standard deviation (SD) among all similar DUs. Use this SD to recalculate Chebyshev UCLs for the DU. i. If the Chebyshev UCL exceeds the decision threshold, follow Line #1 directly above. ii. If the Chebyshev UCL calculated with the worst case SD does not exceed the decision threshold, no additional sampling is needed. d. Because the SU-ICS samples do not represent estimates of the DU mean, the variability between SU-ICSs can be expected to be higher than the variability between DU-ICSs, which will increase the distance between the calculated mean and the UCL. e. Again, because the SU-ICS samples are not estimates of the DU mean, the t-UCL cannot be used unless there are enough SU-ICSs to establish that the distribution of SU-ICS results is normally distributed. Therefore, a 95% Chebyshev (nonparametric) UCL must be used. Chebyshev UCLs are higher than corresponding t-UCLs. f. At least 3 SU-ICSs are needed to calculate a reliable UCL for the DU. g. UCL calculations can be performed, or explored, using the Excel "ICS-95UCL calculator" which is programmed with the following UCL equations.	
5.	Revised CSM page 3 paragraph 1 paragraph 2	Infiltration of storm water through the cap or the base of the ditches will potentially mobilize residual contamination in the sink hole. Based on current flow data, the amount of infiltration occurring between the sink hole and New Cricket Spring may be negligible, but in wetter years that may not be	Dye tracing continues to be the most appropriate technology for the hydrogeologic setting at the site. Water moves downward through the clean cover and weathered residuum (the original topsoil was removed in constructing the site) and into the

	Figure 6	true. Either colloidal transport or dissolution or both may be occurring and either may be the primary mechanism for dioxin movement in groundwater at this site. New technology is available to better assess this movement and the concentration reaching offsite wells or springs, and different standards now apply. The groundwater transport pathway should be considered complete, and additional decision unit(s) should be added to assess impacts to off-site receptors.	epikarstic zone. This is the weathered upper part of the limestone bedrock. Rock solution has developed localized and hydrologically integrated lateral flow routes through this zone. The epikarstic zone hydrologically functions to detain and transport water laterally to drainage features. New Cricket Spring is the primary (and probably the sole) drainage feature for the vicinity of the former onsite sinkhole. The former onsite sinkhole, like many sinkholes, was developed into the epikarstic zone. The planned dye trace will verify and characterize flow from the vicinity of the former sinkhole to New Cricket Spring. The proposed dye tracing is the most appropriate test for assessing and characterizing water movement from the vicinity of the sinkhole to New Cricket Spring. Water from New Cricket Spring is being treated to remove contaminants of concern. As a result, the groundwater contaminant pathway from the main portion of the site flows toward the treatment plant at New Cricket Spring. The potential for contaminants to migrate to offsite wells was evaluated during the 1991 dye tracing study. During the 1991 tracing program the Leatherman and Birmingham Wells were monitored for tracer dyes without any dye detections. Public water supply lines have subsequently been extended to nearby residents. See attached letter dated April 19, 2004.

6.	Revised CSM page 3 paragraph 1 paragraph 2 Figure 6	As stated above, the groundwater transport pathway should be considered complete. The off-site residential receptor should be included in the Conceptual Site Model. Please revise the text and figure.	
7.	Revised CSM page 3 paragraph 1 paragraph 2 Figure 6	All potential receptors should be evaluated, as the McKesson Corporation cannot control the off-site usage. If results of updated Tracer Study identify additional offsite wells or springs that receive dye, then those locations should be evaluated for dioxin as well. See similar comments on Tracer Study Plan (including Tracer Study Comment No. 4, a general comment on that plan).	
8.	Revised CSM page 5 Decision Unit No. 1	The CSM states that "no treated wood storage or processing activities were conducted based on available information." However, an aerial photo from 1970 is available that shows the same activities occurring in this area as in the main area of the site.  Aerial photo, April 9, 1970 It is not clear whether this area was ever sampled at all. If it is contaminated in the ppb range (1,000 times the ppt range of	

		the screening level), then sampling just 2 SUs will be enough to establish this. If the area has concentrations in the ppt range, ProUCL or a similar statistical software package can be used to estimate concentration results that will produce a Chebyshev UCL below 730 ppt with only 2 SUs.	
9.	Revised CSM page 5 Paragraph 1 ROD p. 65	This paragraph says "*All of the proposed samples will be surface soil samples collected from 0-2 inches in depth.*" However, the remedy described in the ROD calls for the entire site to be covered with 6 to 12 inches of clean topsoil. Thus, the proposed samples should be collected from a minimum of 0-6 inches in depth.	
10.	Revised CSM page 5 Decision Unit No. 2	Under most situations, EPA would recommend no less than 8 SUs so that the statistical distribution of the SU data could be determined, and it would not be necessary to default to a nonparametric UCL (which are higher than parametric UCLs). However, if the concentrations are as low as McKesson believes, then the nonparametric Chebyshev UCL (that would have to be used with 5 SU results) will probably not be a problem. However, it is useful to explore what could happen statistically with different types of data sets. Doing this can help refine a sampling design so that the chance of needing to come back and collect more samples can be balanced against the cost/benefit of collecting more samples in the first go. For example, as an upper bound on potential data outcomes, data exploration shows DU#2 **DU ID:** Exploration Data pt 1 580.0 Data pt 2 550.0	

		that for an n = 5, a mean = 615.6, and a SD = 58.7, the 95% Chebyshev UCL = 730.0 ppt. Using statistics to estimate the DU mean for comparison to the screening level of 730 ppt has the following 2 ramifications: • Even if all of the SU-ICS sample concentrations are less than 730 ppt, but the variability in the data set is such that the UCL exceeds 730 ppt, the decision is that, at 95% decision confidence the DU average may exceed the screening level. There are then 2 options: assume the DU is "dirty" and proceed Data pt 3 600.0 Data pt 4 698.0 Data pt 5 650.0 Data pt 6 Data pt 7 Data pt 8 Data pt 9 Data pt 10 Mean = 615.6 Total Std Dev = 58.7 n = 5 1-sided 95%t-UCL = 671.6 RSD (as %) = 9.5 Chebyshev 95UCL = 730.0	
		UCL = 731.3 475.3 619.0 RSD (as %) = 0.6 173.6 15.3 Chebysh ev 95UCL = 735.7 784.9 688.6 accordingly, or collect additional data that can bring the UCL under the screening level. • Even if 1 or more SU-ICS sample concentrations EXCEED the 730 ppt screening level, if the 95% UCL	

		is LESS than the screening level, the decision is that there is 95% confidence that the DU <u>average</u> does NOT exceed the screening level. However, this is not easy to achieve with 5 samples as the following snapshot from the UCL calculator illustrates: **DU ID:** DU#2 Exploration / DU#2 Exploration / DU#2 Exploration Data pt 1: 725.0 / 40.0 / 450.0 Data pt 2: 725.0 / 40.0 / 475.0 Data pt 3: 725.0 / 40.0 / 500.0 Data pt 4: 725.0 / 40.0 / 525.0 Data pt 5: 735.0 / 735.0 / 550.0 Data pt 6: 575.0 Data pt 7: 600.0 Data pt 8: 625.0 Data pt 9: 650.0 Data pt 10: 735.0 Mean = 727.0 / 179.0 / 568.5 Total Std Dev = 4.5 / 310.8 / 87.1 n = 5 / 5 / 10 1-sided 95%t- Each of the data columns has 1 value that slightly exceeds 730 (red frame). The first data column (the farthest left) minimizes variability by having all results close together (but 4 of them below 730), but the concentration will always be over 730 (blue frame). The second data column shows how the Chebyshev UCL will exceed 730 even if all other results are very low such that the mean is very low. The UCL is high due to the high variability created by a single high result. On the other hand, when 10 SU results are available, 1	

		or more individual SUs could exceed the screening level without pushing the UCL over, as long as the other results were low enough and consistent enough for the mean and SD to be low. In addition, since there are 10 data points, ProUCL can be used to test the data set's distribution. Since the third data set is normally distributed, the t-UCL (green frame, 619 ppt) would be appropriate to use.	
11.	Revised CSM page 5 Decision Units No.3 and No. 4	EPA accepts the proposal for sampling of the 2 storm water ditches.	
12.	Revised CSM page 5 Decision Unit No. 5	EPA does not agree with Table 2's expectation that heterogeneity in this area is "moderate." The 2012 data, as mentioned in the proposed plan, were 328 and 1600 ppt TEQ. If those results are put into the statistical software, the mean is 964, and the SD = 900. This produces a Chebyshev UCL of nearly 4000 ppt. The big unknown is whether "most" of the potential samples in this area would have concentrations closer to 300 or 1,600 ppt. If the 1,600 ppt is thought to be an anomaly, the hope might be that the average concentration for this 720 sq.ft.is below the 730 ppt screening level. However, given what is known already about this area's heterogeneity, at least 3 replicate DU-ICSs would be required to measure the field variability and calculate a reliable UCL. Because the area is so small, EPA would agree to fewer increments (e.g., 20) per ICS sample. But if the goal of sampling is to show that the concentration is less than 730, **at least 3 DU-ICS replicates are needed.** On the other hand, if the thought is that a single DU-ICS sample will exceed, and some cleanup activity will be required,	

		then EPA could accept a single ICS result and no UCL. If the DU-ICS result exceeds 730 ppt, then the sampling design worked out. However, if the DU-ICS sample comes back less than 730 (even if it were only 50 ppt), then more work would be required, as discussed in the paragraph above, to establish that the DU mean (as estimated by a UCL) is less than 730 ppt.	
13.	Revised CSM page 6 Decision Unit No. 6	Because of the higher level of heterogeneity near the wash pad, the ¼-acre SU for the wash pad area should have 3 DU-ICS replicates. EPA agrees that the other SU may have one 30-increment DU-ICS. The average (not a UCL) of the 3 replicates SU and the single SU-ICS result **(n = 2) would be used to calculate a UCL for the DU**. The issue of a very low n, along with potentially large variability between the 2 SU results increases the chance that the UCL will exceed 730 ppt even if the calculated mean is fairly low. This could necessitate returning to the site to resample if demonstrating that the DU is "clean" is the expected goal. If the UCL exceeds 730, and it looks like only the wash pad is "dirty," and it is desirable to not clean up the rest of the DU, at least 1 more SU will need to be sampled so that there will be an n of at least 2 to calculate the new DU's (3 SUs, without the wash pad) UCL.	
14.	Revised CSM page 6 Decision Unit No. 7	EPA concurs with 1 DU-ICS with 30 increments for this small area expected to have low variability.	
15.	Revised CSM Figure 6 ROD Page 22, 52 Corrected Deed	Figure 6 provides evaluation of risk to a maintenance worker. However, the ROD identifies an industrial worker as an acceptable receptor at this site and it says that site access will be limited, and the Deed Notice restricts subsurface digging or disturbances. Also, the non-carcinogenic industrial soil screening level is calculated based upon a standard worker scenario. To be consistent with the terminology used in the	

	Notice and Restrictions Page 3	ROD, please identify the industrial worker as a receptor at the site. This would also address the possibility of industrial workers at nearby affected locations. EPA acknowledges that exposure to the surface soil (0 to 6 inches) is the only complete exposure pathway at the site, similar to the maintenance worker. Please adjust the figure to read industrial worker to align with future land use and risk screening tables.	
16.	Anticipated Future Use of Adjacent Property	Regarding: (1) the potential sale of 12 acres adjacent to the southeast end of the Arkwood site, (2) other properties nearby where dye may show up, and (3) the effects of such events on the revised CSM and DUs, it will be appropriate to evaluate such subjects following completion of the supplementary groundwater dye-tracing investigation. This will enable consideration of the intended reuse scenario in the context of a revised CSM. If the tracer study shows water flows from the site to other properties, then that would indicate additional pathways exist and dioxin sampling would be needed.	
18.	Anticipated Future Use of Adjacent Property	Access to the adjacent 12 acres must be kept separate from the rest of the National Priorities List (NPL) site.	

Saved as S:\Tom\ Arkwood32 CSM consolidated comments_20140718 final.doc